STAT 201B: Probability

Feynman Liang*
Department of Statistics, UC Berkeley

Last updated: February 11, 2020

Contents

1 Lecture 6

2020-02-6

2 Lecture 7

2020-02-11

Useful facts should know by heart: X_i IID given $N = n$, $S_N = \sum_1^N X_i$, $N \in \mathbb{N}^*$ random, then if X_i are $\mathbb{N}^*$-valued and $\phi_X(z) = \mathbb{E}z^X$ its generating function,

$$\phi_{S_N}(z) = \phi_N(\phi_X(z)) = \mathbb{E}[\phi_X(z)^N]$$

because sums of RVS becomes products of MGFs.

We saw this in the homework where $X \sim \text{Bern}(p)$ (so $\phi_X(z) = q + pz$) and $N \sim \text{Pois}(\mu)$ gives

$$\phi_{S_N}(z) = e^{\mu(q+pz-1)} = e^{\mu p(z_1)}$$

So a sum of a Pois(μ) number of IID Bern(p) variables is Pois(μp). This is *Poissonization of binomial*, the first step of *Poissonization of multinomial*.

2.1 Limit theorems for Markov chains

- Transition probabilities
- Stationary distributions
- Ergodic theoremm

Let S be a countable state space, P a fixed transition matrix (row-stochastic) on S, $x, y \in S$ states. Assume P is irreducible, i.e. single communication class.

Either all states are transient, which occurs iff $G = \sum_n^\infty P^n$ has finite entries. $G(x, y) = \mathbb{E}_x N_y = \mathbb{E}_x \sum_0^\infty \mathbb{1}\{X_n = y\}$. Or, all states are recurrent $\Leftrightarrow G(x, y) = \infty$.

Assume P irreducible recurrent (so all x are recurrent). Look at iterates of P, P^n being the n step transition matrix. Ask: what happens to $P^n(x, y)$ as $n \to \infty$?

*feynman@berkeley.edu

2.1.1 Stationary measures

Nice to allow globally infinite, locally finite $\mu = (\mu(x), x \in S)$. μ is a probability row vector, and we write

$$(\mu P)(y) = \sum_x \mu(x)P(x,y)$$

Say μ is *invariant measure* if $0 \le \mu(x) < \infty$ for all $x \in S$ and $\mu P = P$.

The existence of a stationary μ tells you very little.

Example 1

Nearest neighbor RW on $\mathbb{Z}$, $p+q=1$. We know $p \neq q \Rightarrow$ transient (LLN) and $p = q \Rightarrow$ recurrent. Also $P^{2n}(0,0) = \binom{2n}{n}2^{-2n} \sim c/\sqrt{n}$.
But, if there exists a finite stationary distribution, then P is recurrent.
Proof: Kac identity (TODO). Consider MC (π, P) with $X_0, X_1 \to \pi(y)$. This is a strictly stationary process

$$(X_1, X_2, \ldots) \overset{d}{=} (X_0, X-1, \ldots)$$

Apply Kac identity to 0/1 process $\mathbb{1}\{X_n = x\}$ for some x to conclude x is recurrent.

Now go the other way.

Theorem 2

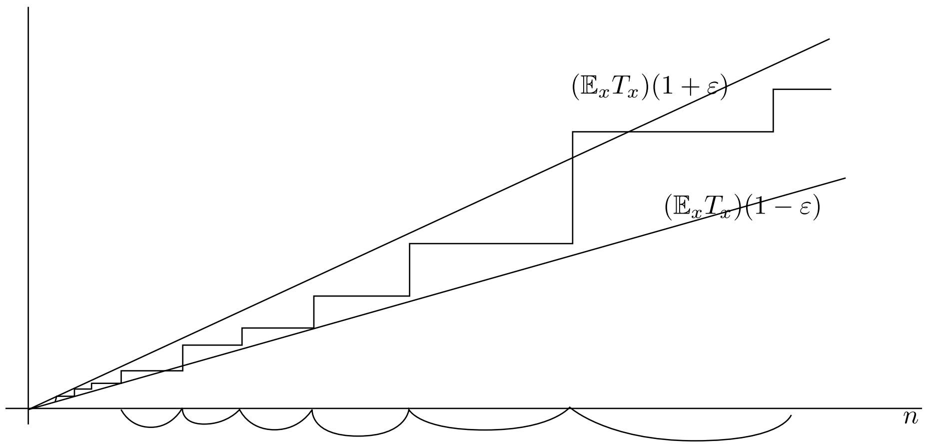
Assume P is ?? on S. TFAE

- *$\exists$ a stationary probability vector π*
- *For some $x \in S$, $\mathbb{E}_x T_x = \mathbb{E}_x \min\{n \ge 1 : X_n = x\} < \infty$*
- *For all $x \in S< \mathbb{E}_x T_x < \infty$*

When all these hold (and P is recurrent), $\pi_x = (\mathbb{E}_x T_x)^{-1}$ is unique. Also, the expected number of hits on y before you come back

$$\mathbb{E}_x \sum_{n=0}^{\infty} \mathbb{1}\{X_n = y, T_x > n\} = \textit{mean number of ys in an x block} = \pi_y/\pi_x$$

Let $T_x^{(n)}$ denote the time of the nth return to x. This is a sum of IID copies of T_x. $T_x^{(n)}/n \overset{as}{\to} \mathbb{E}_x T_x$. $\Rightarrow \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}\{X_n = x\} \overset{as}{\to} \frac{1}{\mathbb{E}_x T_x}$



{fig:as-in-w
edge} Figure 1: Eventually it will stay between any $\pm\varepsilon$ wedge.

The $\mathbb{E}_x T_x$ is from the LLN, the $\cdot^{-1}$ from inverting.

This is the basics of *ergodic theory of MCs*, driven by some extension of SLLN. APply SLLN to x-cycles to learn in the long run if $\mathbb{E}_x T_x < \infty$ then the expected number of visits to x within unit time converges to $(\mathbb{E}_x T_x)^{-1}$. Also, letting $\mu_x(y)$ be the mean number of ys in a x block we have that $\mu_x(y) < \infty$ even without $\mathbb{E}_x T_x < \infty$.

Notice the *fundamental relation*

$$\mathbb{E}_x T_x = \sum_{y \in S} \mu_x(y)$$

Also, can apply SLLN to y to conclude in the long run

Notice, we get $\mu_x(y) = \frac{1}{\mu_y(x)}$ by LLN. Hence, it's plausible that $\mu_x(y) = \frac{\mu(y)}{\mu(x)}$ for some (hence every) stationary measure μ. Also, if $\mathbb{E}_x T_x < \infty$ then $\sum_y \mu_x(y) = \mathbb{E}_x T_x$ which implies

$$\frac{\mu_x(x)}{\mathbb{E}_x T_x} = \pi_x = ?? \text{ prob of } x$$

because $\mu_x(x) = 1$ by definition. This is motivation for guessing that $y\mu_x(y)$ is the unique invariant probability measure with mass 1 at x. THis is true for any recurrent chain: for a positive recurrent chaing $\mathbb{E}_x T_x < \infty \Leftrightarrow \sum_y \mu_x(y) < \infty$ which implies $\pi_x = (\mathbb{E}_x T_x)^{-1}$.

- Easy: $\sum_y \mu_x(y) = \mathbb{E}_x T_x$
- (to be shown) $y \mapsto \mu_x(y)$ is stationary: $\mu_x(\cdot)P = \mu_x(\cdot)$ and $\mu_x(x) = 1$.

A stationary *probability* measure, if it exists, is unique. Renormalize $\mu_x(y)$ by $\mathbb{E}_x T_x$ so $\pi_y = \frac{\mu_x(y)}{\mathbb{E}_x T_x}$, then $\sum_y \pi_y = 1$.

Sketch of why $\mu_x(\cdot)P = \mu_x(\cdot)$. Basic idea is to sum a geometric progression:

$$(I + P + \cdots + P^{n-1}) + P^n = I + (I + P + \cdots + P^{n-1})P$$

Probability interprpetation: take an initial distribution λ.

$$\lambda P^n = \Pr_\lambda(X_n \in \cdot)$$

$$\lambda(I + P + \cdots + P^{n-1}) = \sum_{k=0}^{n-1} \lambda P^k(\cdot) = \mathbb{E}_\lambda \text{num hits on } \cdot \text{ in times } [0, n-1]$$

Now we do this with n a stopping time T.

Claim: Let (X_n) be a time-homogeneous MC with TM P, $X_0 \sim \lambda$, $\Pr_x(T < \infty) = 1$, and

$$\Pr_\lambda(X_T \in \dot{)} = \lambda_T(\cdot) = \text{distn of } X_T$$

Define the Green measure

$$\lambda G_T(\cdot) \coloneqq \mathbb{E}_\lambda \sum_{n=0}^{T-1} \mathbb{1}\{X_n \in \cdot\}$$

Then

$$\lambda G_T + \lambda_T = \lambda + (\lambda G_T)P$$

This works when T is a fixed time, and the claim is that it holds for a stopping time T. This is called an *occupation measure identity*, and notice if $\lambda_T = \lambda$ we have that the occupation measure λG_T is stationary.